

UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF NEW YORK

----- X
:
:
In re OpenAI, Inc. Copyright Infringement :
Litigation : No. 1:25-md-03143-SHS-OTW
:
:
:
----- X

STIPULATION & ORDER RE: TRAINING DATA INSPECTION PROTOCOL

Upon the stipulation of the parties, the following protocol will apply to the inspection, review, and/or disclosure of Training Data produced by Defendants OpenAI, Inc., OpenAI, L.P., OpenAI L.L.C., OpenAI OpCo, L.L.C., OpenAI Global L.L.C., OAI Corporation L.L.C., OpenAI Holdings L.L.C., OpenAI GP, L.L.C., OpenAI Startup Fund GP I, L.L.C., OpenAI Startup Fund I, L.P., and OpenAI Startup Fund Management, LLC (collectively, “OpenAI”):¹

1. For the purposes of this protocol, “Training Data” shall be defined as data used to train relevant OpenAI LLMs, though the parties may, upon mutual agreement, use this protocol for other data as well (e.g. data regarding output filtration). OpenAI reserves the right to update the Training Data made available for inspection should it find additional responsive data, and Plaintiffs reserve the right to request additional Training Data that has been disclosed or discovered and is not made available for inspection.

2. The “Inspecting Party” shall be defined as any of the Plaintiffs to *Tremblay, et al. v. OpenAI, Inc., et al.*, No. 3:23-cv-03233 (N.D. Cal.), *Silverman, et al. v. OpenAI, Inc., et al.*, No. 3:23-cv-03416 (N.D. Cal.), *Chabon, et al. v. OpenAI, Inc., et al.*, No. 3:23-cv-04625 (N.D. Cal.), *In re OpenAI ChatGPT Litigation*, No. 3:23-cv-0322 (N.D. Cal.), *Millette v. OpenAI, Inc., et al.*, No. 5:24-cv-04710 (N.D. Cal.), *Authors Guild, et al. v. OpenAI, Inc., et al.*, No. 1:23-cv-08292 (S.D.N.Y.), *Alter, et al. v. OpenAI, Inc., et al.*, No. 1:23-cv-10211 (S.D.N.Y.), *The New York Times Co. v. Microsoft Corp., et al.*, No. 1:23-cv-11195 (S.D.N.Y.), *Basbanes, et al. v. Microsoft Corp., et al.*, No. 1:24-cv-01514 (S.D.N.Y.), *Raw Story Media, Inc., et al. v. OpenAI, Inc., et al.*, No. 1:24-cv-01514 (S.D.N.Y.), *The Intercept Media, Inc. v. OpenAI, Inc., et al.*, No. 1:24-cv-01515 (S.D.N.Y.), *Daily News LP, et al. v. Microsoft Corp., et al.*, No. 1:24-cv-03285 (S.D.N.Y.), *The Center for Investigative Reporting, Inc. v. OpenAI, Inc., et al.*, No. 1:24-cv-04872 (S.D.N.Y.), the

¹ By entering into this stipulation, Plaintiffs do not intend to waive any rights they may (or may not) have to seek discovery of Training Data by means other than by inspection.

Ziff Davis, Inc. et al. v. OpenAI Inc., et al., No. 1:25-cv-4315 (S.D.N.Y.) litigation, and any actions subsequently transferred to this Court pursuant to Rule 7.1 of the Rules of Procedure of the Judicial Panel on Multidistrict Litigation or otherwise centralized with this litigation (centralized as *In re: OpenAI, Inc. Copyright Litigation*, No. 1:25-md-03143 (S.D.N.Y.) and collectively the “Centralized Actions”, and each individual action, an “Action”), including their attorneys of record, agents, retained consultants, experts, and any other persons or organization over which they have direct control.

3. Training Data shall be made available for inspection in electronic format at OpenAI’s offices in San Francisco CA, or at a secure location determined by OpenAI within 25 miles of San Francisco, CA; or at another mutually agreed location. Training Data will be made available for inspection between the hours of 8:30 a.m. and 5:00 p.m. on business days, although the parties will be reasonable in accommodating reasonable requests to conduct inspections at other times. With prior notice from the Inspecting Party, OpenAI shall make a reasonable effort to ensure that the Training Data is programmatically accessible on a continuous, twenty-four-hour basis across multiple days as needed in order to permit the Inspecting Party to run reasonable searches or other automated programs to analyze the Training Data.

4. The Inspecting Party shall provide five (5) calendar days’ notice prior to any inspection.

5. Training Data shall be designated “HIGHLY CONFIDENTIAL – ATTORNEYS’ EYES ONLY” pursuant to the Stipulated Protective Order entered in the Centralized Action on July 24, 2025, and the Inspecting Party may disclose Training Data only to those authorized to view “HIGHLY CONFIDENTIAL – ATTORNEYS EYES ONLY” information under paragraph 24 of the Stipulated Protective Order, without prejudice to any party’s right to challenge this confidentiality designation (or oppose a challenge to the confidentiality designation) at a later date.

Any challenge to the confidentiality designation of the Training Data or portions thereof under this Training Data Inspection Protocol shall be written, shall be served on outside counsel for OpenAI, shall particularly identify the documents or information that the Inspecting Party contends should be differently designated, and shall state the grounds for the objection. The parties shall meet and confer in a good faith effort to resolve the dispute. Notwithstanding any challenge to a designation, the Training Data in question shall continue to be treated as “HIGHLY CONFIDENTIAL – ATTORNEYS’ EYES ONLY” until one of the following occurs: (1) OpenAI withdraws such designation in writing; or (2) the Court rules that the Training Data in question is not entitled to the designation.

6. Nothing in this Training Data Inspection Protocol shall alter or change in any way the requirements of the Stipulated Protective Order. In the event of any conflict, however, this Training Data Inspection Protocol shall control for any Training Data made available for inspection.

7. Training Data shall be made available for inspection and review subject to the following provisions:

a. Training Data shall be produced as maintained by OpenAI in the ordinary course of business and made available by OpenAI in a secure room on a secured computer with necessary network access to a host computer containing the Training Data but without Internet access or network access to other unauthorized computers or devices (together the “secured computer”). The secured computer will contain a README file that will provide a directory of the Training Data and brief descriptions of layout, format, and searching, which will be produced to the Inspecting Party in advance of any inspection.

b. The secured computer will be equipped with tools that are sufficient for viewing and searching the Training Data made available for inspection. Any data made available

for inspection will be in its native format as it is used during the normal course of business.

Defendants will reasonably cooperate with Plaintiffs to address any technical concerns Plaintiffs may have regarding the form of production of the Training Data and the hardware and software that is provided so that Plaintiffs may conduct the Training Data review pursuant to the applicable rules of the Federal Rules of Civil Procedure. Plaintiffs reserve all rights to seek any additional relief from the Court, including to enable a more efficient and/or effective review of the Training Data.

c. The Producing Party shall provide the Inspecting Party with information explaining how to start, log on to, and operate the secured computer in order to access the Training Data on the secured computer. An individual will be available onsite to handle technical support issues, and the Producing Party's outside counsel will be available electronically to make reasonable efforts to attempt to resolve issues that may arise during the course of inspection.

d. The Inspecting Party's counsel and/or experts may request that software tools and/or files be installed on the secured computer, provided, however, that (a) the Inspecting Party possesses an appropriate license to such software tools and/or files; (b) OpenAI approves such software tools and/or files, such approval not to be unreasonably withheld; and (c) such other software tools and/or files are reasonably necessary for the Inspecting Party to perform its review of the Training Data consistent with all of the protections herein. The Inspecting Party must provide OpenAI with the licensed software tool(s) and/or files, at the Inspecting Party's expense, at least seven calendar days in advance of the date upon which the Inspecting Party wishes to have the additional software tools and/or files available for use on the secured computer. The Producing Party will install and confirm installation of said software on the secured computer prior to the inspection. OpenAI will reasonably cooperate with the Inspecting Party to accommodate requests to install on the secured computer additional software tool(s) and/or files provided less than seven days in advance.

e. No recordable media or recordable devices, including without limitation computers, cellular telephones, cameras, other recording devices, or drives of any kind, shall be permitted into the secure inspection room, except the Producing Party may provide a limited-use note-taking computer at Inspecting Party's request, solely for note-taking purposes. At the end of each day of inspection, the Inspecting Party shall be able to copy notes from the note-taking computer onto a recordable device, under the supervision of the Producing Party. For the avoidance of doubt, medical devices (e.g., heart monitors, insulin pumps), stopwatches, and timers are permitted so long as such devices are not capable of recording or copying data from the secured computer.

f. The Inspecting Party's counsel and/or experts may take handwritten notes or electronic notes on the provided note-taking computer in scratch files, but may not copy any Training Data itself into any notes. For the avoidance of doubt, this provision shall not prevent the Inspecting Party's counsel and/or experts from recording in their notes statistical information (such as hits, file sizes, or match scores) or particular items, files, or categories of items or files contained in the Training Data. The Inspecting Party will not waive any applicable work-product protection over their electronic notes by saving them to the note-taking computer temporarily. Such notes may not be in encoded or encrypted form. Any notes related to the Training Data will be treated as "HIGHLY CONFIDENTIAL – ATTORNEYS' EYES ONLY."

g. The Producing Party may visually monitor the activities of the Inspecting Party's representatives during any inspection, but only to ensure that there is no unauthorized recording, copying, or transmission of the Training Data (and not to deliberately attempt to read notes or other written material created by the Inspecting Party). Any monitoring must be conducted from outside the room where the inspection is taking place. Observation must be through glass, such as a glass door, windows, or glass walls, allowing reviewers to converse without being audibly

monitored. The Producing Party will make an unmonitored breakout room with Internet access reasonably convenient to and near the secured inspection room available for use by the Inspecting Party.

h. No copies of all or any portion of the Training Data, or other written or electronic record of the Training Data, may leave the secured room in which the Training Data is inspected except as provided herein. The Inspecting Party may obtain print outs of reasonable and limited portions of the Training Data or electronic notes taken on the secure computer to prepare court filings or pleadings or other papers (including a testifying expert's expert report) by following the procedures provided herein. For purposes of this protocol, references to "print," "printing," or "print outs" are understood to refer to a Bates-stamped electronic production (as described in this Paragraph). To make a request, the Inspecting Party shall create a directory entitled "Print Request" and save the desired limited portions of the Training Data or notes in that directory. The beginning of each portion of Training Data the Inspecting Party wishes to print must include the filename, file path, and line numbers where the material was found in the training data or other information that allows for specific identification of the material. The Inspecting Party shall alert OpenAI when it has saved the desired limited portions of the Training Data or notes in the "Print Request" directory that it requests to be printed. Upon receiving a request, OpenAI shall Bates number, and label "HIGHLY CONFIDENTIAL – ATTORNEYS' EYES ONLY" all requested pages. Within 7 business days from the date of request, OpenAI shall either (i) produce electronic versions to the Inspecting Party's counsel, or (ii) inform the Inspecting Party that OpenAI objects that the requested portions are excessive, not for a permitted purpose, and/or not justified (*see, e.g.*, Fed. R. Civ. P. 26(b)). In the event that OpenAI objects, the parties shall meet and confer within three business days of OpenAI's notice of its objection. If, after meeting and conferring, OpenAI and the Inspecting Party cannot resolve the objection, the Inspecting Party shall be entitled to seek a Court

resolution of whether the requested Training Data should be produced. To the extent the Inspecting Party has a right to seek production, separate and apart from any inspection, of portions of the Producing Party's Training Data, nothing in this protocol should be read to prejudice that right.

i. All persons who will review OpenAI's Training Data on behalf of an Inspecting Party, including the Inspecting Party's counsel, must qualify under paragraph 24 of the Stipulated Protective Order as an individual to whom "HIGHLY CONFIDENTIAL – ATTORNEYS' EYES ONLY" information may be disclosed, and must sign the Non-Disclosure Agreement attached as Exhibit A to the Stipulated Protective Order. All persons who review OpenAI's Training Data in the secured inspection room or on the secured computer on behalf of an Inspecting Party shall also be identified in writing to OpenAI at least five business days in advance of the first time that such person reviews such Training Data. All authorized persons viewing Training Data in the secured inspection room or on the secured computer shall, on each day they view Training Data, sign a log that will include the names of persons who enter the locked room to view the Training Data and when they enter and depart. Proper identification of all authorized persons shall be provided prior to any access to the secure inspection room or the secured computer containing Training Data. Proper identification requires showing, at a minimum, a photo identification card sanctioned by the government of any State of the United States, by the government of the United States, or by the nation state of the authorized person's current citizenship. Access to the secure inspection room or the secured computer may be denied, at the discretion of OpenAI, to any individual who fails to provide proper identification.

j. Unless otherwise agreed in advance by the parties in writing, following each day on which inspection is done under this protocol, the Inspecting Party's counsel and/or experts shall remove all notes, documents, and all other materials from the secure inspection room. OpenAI shall not be responsible for any items left in the room following each inspection session, and the

Inspecting Party shall have no expectation of confidentiality for any items left in the room following each inspection session without a prior agreement to that effect.

k. Other than as provided above, the Inspecting Party will not copy, remove, or otherwise transfer any Training Data from the secured computer including, without limitation, copying, removing, or transferring the Training Data onto any recordable media or recordable device. The Inspecting Party will not transmit any Training Data in any way from OpenAI's facilities.

8. Notwithstanding any provisions of this Training Data Protocol or the Stipulated Protective Order the Parties reserve the right to amend this protocol either by written agreement or Order of the Court upon showing of good cause.

IT IS SO STIPULATED, through Counsel of Record.

Dated: August 21, 2025

/s/ Ian B. Crosby

Ian B. Crosby (*pro hac vice*)
SUSMAN GODFREY L.L.P.
401 Union Street, Suite 3000
Seattle, WA 98101
Telephone: (206) 516-3880
icrosby@susmangodfrey.com

Attorney for Plaintiff
The New York Times Company and
News Plaintiffs' Liaison Counsel

/s/ Jennifer B. Maisel

Jennifer B. Maisel
ROTHWELL, FIGG, ERNST &
MANBECK, P.C.
901 New York Avenue, N.W., Suite 900 East
Washington, DC 20001
Telephone: (202) 783-6040
Facsimile: (202) 783-6031
jmaisel@rothwellfigg.com

Attorney for Plaintiffs
Daily News, LP; The Chicago Tribune
Company, LLC; Orlando Sentinel
Communications Company, LLC; Sun-
Sentinel Company, LLC; San Jose Mercury-
News, LLC; DP Media Network, LLC; ORB
Publishing, LLC; and Northwest Publications,
LLC

/s/ Stephen Stich Match

Stephen Stich Match (*pro hac vice*)
LOEY & LOEY
311 North Aberdeen, 3rd Floor
Chicago, IL 60607
312-243-5900 (p)
312-243-5902 (f)
match@loevy.com

Attorney for Plaintiff
The Center for Investigative Reporting, Inc.
and Plaintiff The Intercept Media, Inc.

/s/ Lacy H. Koonce, III

Lacy H. (“Lance”) Koonce, III
KLARIS LAW PLLC
161 Water Street
New York, NY 10038
Telephone: (646) 779-4882
lance.koonce@klaris.com

Attorney for Plaintiffs
Ziff Davis, Inc., Ziff Davis, LLC, IGN
Entertainment, Inc., Everyday Health Media,
LLC, Mashable, Inc. and CNET Media, Inc.

/s/ Justin A. Nelson

Justin A. Nelson (*pro hac vice*)
SUSMAN GODFREY L.L.P.
1000 Louisiana Street, Suite 5100
Houston, TX 77002
Telephone: 713-651-9366
jnelson@susmangodfrey.com

Interim Lead Counsel for Class Plaintiffs

Dated: August 21, 2025

/s/ Thomas E. Gorman

Thomas E. Gorman (*pro hac vice*)
KEKER, VAN NEST & PETERS LLP
633 Battery Street
San Francisco, CA 94111-1809
Telephone: (415) 391-5400
Facsimile: (415) 391-7188
tgorman@keker.com

/s/ Rose S. Lee

roselee@mofo.com (*pro hac vice*)
MORRISON & FOERSTER LLP
425 Market Street
San Francisco, California 94105-2482
Telephone: (650) 813-5865
Facsimile: (213) 327-5074
roselee@mofo.com

/s/ Allison S. Blanco

Allison S. Blanco (*pro hac vice*)
LATHAM & WATKINS LLP
650 Town Center Drive, Ste 2000
Costa Mesa, CA 92626
allison.blanco@lw.com

Counsel for Defendants

OpenAI, Inc.; OpenAI OpCo LLC; OpenAI GP LLC; OpenAI, LLC; OpenAI Global LLC; OAI Corporation, LLC; OpenAI Holdings, LLC; OpenAI Startup Fund I LP; OpenAI Startup Fund GP LLC; and OpenAI Startup Fund Management LLC

IT IS ORDERED that the foregoing Agreement is approved.

Dated: 8/25/2025



HON. ONA T. WANG
United States Magistrate Judge